

Article

Leveraging Machine Learning to Extend Ontology-Driven Geographic Object-Based Image Analysis (O-GEOBIA): A Case Study in Forest-Type Mapping

Sachit Rajbhandari ^{1,*} , Jagannath Aryal ¹ , Jon Osborn ¹ , Arko Lucieer ¹  and Robert Musk ²

¹ Discipline of Geography and Spatial Sciences, School of Technology, Environments and Design, College of Sciences and Engineering, University of Tasmania, Private Bag 76, Hobart, Tasmania 7001, Australia; jagannath.aryal@utas.edu.au (J.A.); jon.osborn@utas.edu.au (J.O.); arko.lucieer@utas.edu.au (A.L.)

² Timberlands Pacific, Level 1, Cimitiere House, 113-115 Cimitiere Street, Launceston, Tasmania 7250, Australia; robert.musk@tppl.com.au

* Correspondence: sachit.rajbhandari@utas.edu.au; Tel.: +61-415-692-370

Received: 11 January 2019; Accepted: 25 February 2019; Published: 1 March 2019



Abstract: Ontology-driven Geographic Object-Based Image Analysis (O-GEOBIA) contributes to the identification of meaningful objects. In fusing data from multiple sensors, the number of feature variables is increased and object identification becomes a challenging task. We propose a methodological contribution that extends feature variable characterisation. This method is illustrated with a case study in forest-type mapping in Tasmania, Australia. Satellite images, airborne LiDAR (Light Detection and Ranging) and expert photo-interpretation data are fused for feature extraction and classification. Two machine learning algorithms, Random Forest and Boruta, are used to identify important and relevant feature variables. A variogram is used to describe textural and spatial features. Different variogram features are used as input for rule-based classifications. The rule-based classifications employ (i) spectral features, (ii) vegetation indices, (iii) LiDAR, and (iv) variogram features, and resulted in overall classification accuracies of 77.06%, 78.90%, 73.39% and 77.06% respectively. Following data fusion, the use of combined feature variables resulted in a higher classification accuracy (81.65%). Using relevant features extracted from the Boruta algorithm, the classification accuracy is further improved (82.57%). The results demonstrate that the use of relevant variogram features together with spectral and LiDAR features resulted in improved classification accuracy.

Keywords: GEOBIA; rule-based classification; ontology; machine learning; random forests; rules extraction; variogram; semantic similarities; semantic variogram

1. Introduction

Geographic Object-Based Image Analysis (GEOBIA) is a widely used and still developing new approach to image segmentation and classification [1]. The goal of GEOBIA is to extract segments, derive meaningful objects, and in turn thematic classes from remotely sensed data. Unlike more traditional approaches to image segmentation and classification, GEOBIA includes contextual information together with a variety of image object properties that include size, shape, texture, and spectral (colour) information. The data used in the segmentation may be internal, extracted from the image and operate by grouping similar pixels into objects, or external to the image and operate by including thematic layers, for example, known land use or other meaningful object boundaries [2]. The aim of any GEOBIA application is to translate expert knowledge associated

with real-world features into the GEOBIA process [1,3,4] in a manner that is formal, objective and transferable. One approach to this challenge is to employ an ontology to formally capture and represent the expert knowledge [4,5], referred to here as Ontology-driven GEOBIA (O-GEOBIA). Ontology helps to reduce semantic gap between high-level knowledge and low-level information. Ontology integrates qualitative (e.g., Forest canopy is dense) and subjective (description referred by subject i.e., experts) high-level knowledge with the quantitative (e.g., spectral band value represented by digital number) and objective (information refers to the object i.e., segmented image object) low-level information [4]. A consequential challenge is to manage decisions regarding the extent to which a prescribed ontology—both the relationships defined by that ontology and any quantitative attributes associated with those relationships—can be treated as transferable or generalisable across different study sites or data sources [6]. In turn, the question that arises is how best to generate the classification rules that cannot be achieved only from the domain knowledge but may be discoverable in the data on a site-specific basis.

An opportunity, but also a further challenge, arises from advances in remote sensing technology that are extending the availability of different types of remotely sensed data, such as multispectral and hyperspectral imagery, radar and LiDAR (Light Detection and Ranging). Thematic geo-information retrieval from these stand-alone datasets would benefit with their fusion. Multi-sensor data fusion techniques are therefore of growing importance [7–11]. In GEOBIA, particularly in forest classification, recent research has shown that 3D LiDAR data can augment imagery data for improved and more robust classification [12] and data fusion can be used to increase the robustness of forest-type mapping [13].

Our research is exploring methodological approaches that integrate ontology into the GEOBIA workflow, with the ontology purposely developed to capture rules that are generalisable and so can be expected to be transferable across different study sites, with these rules supplemented by non-transferable rules developed on a case-by-case basis using data fusion and machine learning. An earlier paper [6] benchmarked how ontological rules, both generalised rules extracted from domain knowledge and localised rules can be incorporated into GEOBIA.

This paper extends that work by developing a methodology for extracting localised rules using machine learning techniques. The methodology develops classification rules using fused multi-sensor data, features extracted from image-based spectral indices and point-cloud derivatives, together with semivariogram features derived within the GEOBIA environment. The methodology employs Boruta algorithm in selecting all-relevant features and includes characterisation of thematic classes using semantic similarities.

Research questions addressed in this paper are:

- Why is multi-sensor data fusion necessary in O-GEOBIA?
- How can spatial features be incorporated for accurate classification in O-GEOBIA?
- How can relevant features be extracted to construct rules required when identifying classes for O-GEOBIA?
- How can semantic similarity be used to characterise thematic classes in an ontological environment?

The novelty of this work lies in its extension of an ontological geographic object-based image analysis framework with respect to data fusion. We aimed to identify the challenges and their remedies in the context of O-GEOBIA. The contributions of this paper are:

- It presents a methodology for improving classification accuracy using feature selection from fused multi-sensor data.
- It evaluates the employment of semivariogram features alongside image-based spectral indices and point-cloud based airborne LiDAR derivatives.
- It presents a methodology for the selection of semantic similarities for semantic characterisation.

The application of our methodology is illustrated using a suitably complex case study: the classification of forest types. The case study demonstrates how existing domain knowledge can be represented by an ontology that is generalisable and how machine learning can be used to supplement the classification with local, case-specific and non-transferable rules. The case study is not intended to establish the robustness or efficacy of our approach, but to illustrate how the method is applied in practice.

The remainder of the paper is organised as follows. Section 2 provides necessary theoretical background. Section 3 outlines the methodology for ontology-based image classification using machine learning algorithms and experimental design. Section 4 presents the case study of Tasmanian forest-type mapping. Section 5 elaborates the implementation of the proposed methodology for the forest-type case study. In Section 6, the results of the case study is presented and discussed in Section 7. In the concluding Section 8, the contribution of the current work is presented.

2. Background

Ontology has been used in the field of image interpretation [4,14–19]. The application of ontology in GIScience for extraction of geo-information has been demonstrated by [20–24]. GEOBIA framework was introduced in 2000 and has proven to be a powerful tool for information extraction from imagery [1,25–27]. There is now an opportunity to incorporate ontological concepts into the GEOBIA framework to improve extraction of meaningful GIS-ready information for further analysis and interpretation [4].

In O-GEOBIA, remote-sensing techniques are used to interpret physical properties (e.g., an NDVI may be used to discriminate vegetated areas from urban) while domain knowledge is used to provide contextual information (e.g., that vegetated areas in an urban setting may be municipal parks) [28]. Low-level information extracted from sensor data in combination with high-level information from domain knowledge provides a basis for creating rule sets for a rule-based classifier.

An ontological framework that can formalise such knowledge for analysis of remote sensing images has been proposed in [6]. An ontology is a formal description of knowledge as a set of concepts and their relationships within a domain. The domain concepts in an ontology represent thematic classes. Gruber [29] states that an ontology is a “specification of a shared conceptualisation”. Ontology is a shared understanding of a domain that formally defines components such as individuals (instances of objects), classes, attributes and relations as well as restrictions, rules and axioms.

Data fusion aims to integrate multi-sensor data to extract information that cannot be derived from the data from any single sensor. In remote sensing, data fusion can take place at three different processing levels: pixel level, feature level and decision level [30]. Pixel level data fusion is a low processing level merging of raw data from multiple sources into common resolution data. An example of pixel level image fusion is pan-sharpening, which aims to improve spatial and spectral resolution along with structural and textural details [11]. Feature level fusion is a high-level fusion that involves the extraction of objects identified in different data sources using segmentation techniques. The alignment of similar objects from multiple sources is performed and various spectral, textural and spatial features are extracted and fused together for statistical or neural network assessment [30]. Decision level fusion merges the extracted information such as selection of a relevant feature from extracted features. Benefits of data fusion embrace the issues of correlated, spurious, or disparate data. Highly correlated data can lead to positively biased results and artificially high confidence levels; spurious data leads to outliers; disparate data leads to conflicting information [7].

In order to describe complex thematic classes and to achieve better classification accuracy, the feature dimension can be further increased with the addition of derivatives computed from existing features. For example, several vegetation indices may be calculated using combinations of different spectral bands. Spectral vegetation indices are commonly used to characterise forests and monitor forest resources [31]. Further, there is a need to understand the spatial dependency of feature variables

for e.g., spatial autocorrelation. In understanding spatial patterns, semivariograms may be used to measure spatial dependencies of feature variables [32].

Increased dimensionality provides new input variables for classification rules, but it also can make it difficult to identify feature variables that play an important role in classifying a particular class. Increasing the feature dimensions adds complexity to segmentation and classification. To tackle the complexity associated with high dimensionality, previous research has proposed the use of filtering and wrapping approaches for feature selection and reduction [33,34]. In our classification work, we adopt a wrapper method because of its strong relationships with the classifier. Wrapper methods are computationally costly compared with filter methods. However, the Boruta algorithm [35] uses a Random Forest (RF) classifier [36] which makes it relatively fast due to its simple heuristic feature selection procedure. In the work reported here, we used Boruta algorithm for feature selection. Boruta is a feature reduction algorithm and it follows an all-relevant variable selection method rather than a minimal-optimal method, taking account of multiple relationships among variables [35]. Boruta runs several Random Forest models to obtain a statistically significant division between relevant and unimportant feature variables. The feature reduction produces a reduced dataset, which can be expected to improve classification accuracy due to the elimination of noise.

Machine learning (ML) has the potential to contribute to improved feature selection and to generate implicit knowledge from the fused data. Rules may be extracted using ML very quickly and extracted rules are often comparable to human-crafted rules [37,38]. Data fusion of multi-sensor data can assist in defining accurate classification rules. In our work, ML is used to develop new rules by extracting them from the data itself and applying those rules in an O-GEOBIA framework.

In this paper, we present an ontology-based approach to determine the similarity between two classes and recommend semantic similarity measures that work for multi-sensor data. Semantic similarity is one approach to quantifying the similarity between two different classes. Semantic similarity measures are widely used in Natural Language Processing [39] and Ontology Alignment [40] and are becoming important components of knowledge-based and information retrieval systems. Ontology-based semantic similarity measures are categorised into hierarchy-based, information content based, and feature-based [39].

In our case study, we implemented feature level data fusion from multispectral RapidEye satellite imagery, airborne LiDAR data, and PI (Photo Interpretation) data. The fusion of multi-sensor data contributed to the extraction of spectral, spatial, and contextual features, which were used to develop classification rules in GEOBIA. The Boruta algorithm was employed to extract all relevant variables in order to define more accurate classification rules. Semantic similarity methods were used to characterise the similarities between different classes and we identified similarity measures that are appropriate for an O-GEOBIA framework.

3. Methodology

In this section, we first present an overall methodological workflow for extended ontology driven object-based image analysis. Next, we explain the contextual experimental design developed for this study.

3.1. Extension of an Ontology-Driven Geographic Object-Based Image Analysis (O-GEOBIA) Framework

This study proposes an extension of an O-GEOBIA framework [6] that uses ML techniques method for automatic generation of rules. Figure 1 shows the methodological steps, which comprises: (1) data pre-processing, (2) feature selection based on ML, (3) rules generation using ML and ontology, (4) ontology based image classification, and (5) semantic characterisation. These 5 steps are categorised into 3 different stages. Our research work is largely focussed on Stage 1 (data fusion and feature selection) and Stage 3 (semantic characterisation). The Stage 2 (rules generation and ontology based image classification) is applied based on our previous work [6].

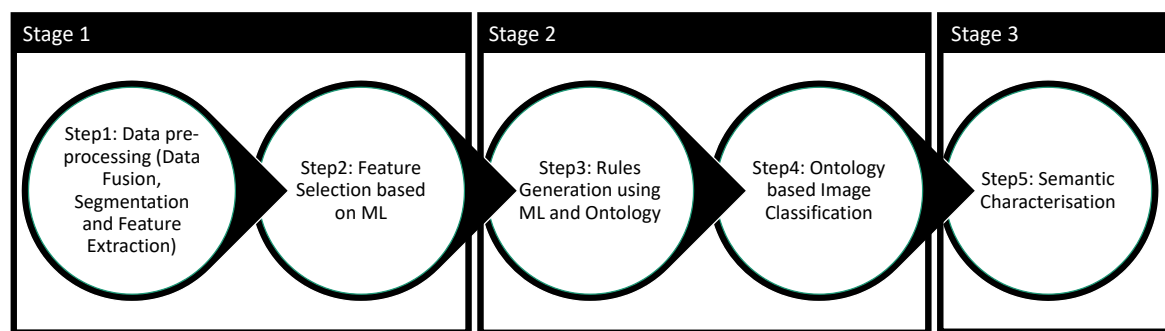


Figure 1. An overall methodological workflow for Ontology-driven Geographic Object-Based Image Analysis (O-GEOBIA).

Step 1 Data pre-processing

This component comprises a fusion of multi-sensor data, image segmentation and feature extraction. Different multi-sensor data such as satellite imagery and LiDAR are fused, resulting in a new group of features. Image segmentation is carried out to delineate image objects. For each image object, their underlying features value is calculated. Depending on the kind of data used for fusion, different feature variables, such as spectral and spatial, are extracted. The output of this module are extracted feature variables from multi-sensor data, which are input for the next module.

Step 2 Feature selection based on ML

With data fusion, a high number of features are available. In this component, we select relevant features using machine learning techniques. To achieve this, we use the Boruta algorithm developed as a wrapper around the Random Forest classifier for identification of important and relevant variables. In this work, we aim to illustrate the importance of feature selection in multi-sensor data with experimental results.

Step 3 Rules generation using ML and Ontology

After selection of features, we use the inTrees (interpretable Trees) framework for automatic extraction of classification rules from the datasets. These rules are added to an ontology along with the expert-defined rules from the next module.

Step 4 Ontology based image classification

For image classification, the ontological framework proposed in [6] is adopted. The classification experiments are based on spectral, LiDAR and variogram based features.

Step 5 Semantic characterisation

Finally, for semantic characterisation, semantic similarities between the different domain classes, as defined in an ontology, are measured. Based on semantic distances, a semantic variogram is calculated for the characterisation of domain classes. The semantic variogram is used as a metric to characterise the variability between classes based on semantic distances.

3.2. Contextual Experimental Design

The experimental design for forest characterisation is carried out using three approaches: (1) feature attributes; (2) spatial relations; and (3) semantic relations:

1. The feature attributes approach helps in the creation of classification rules but ignores spatial relationships. The measurement data used in this approach include spectral and LiDAR data.

2. The spatial relationships approach addresses spatial relationships and specifically contributes to the classification using measures of autocorrelation. A variogram is used in this approach.
3. The semantic relations approach is based on the semantic relations between classes and uses an ontology. A semantic variogram is used for this approach.

3.2.1. Feature Attributes

Different spectral bands from the sensors are used as feature attributes. From the LiDAR data, we extracted various derivatives such as elevation, height and intensity. From the PI data (described in Section 4) we extracted the forest structural group and class (FC2011) information. The feature variables from different data sources are listed in Table 1.

Table 1. Variables extracted from multi-sensor data for classification.

Type	Name
Spectral	Brightness
	Blue
	Green
	NIR
	Red
LiDAR	Red Edge
	Canopy Height (CH)
	Height
	Canopy Intensity (CI)
	Intensity
	Elevation
	Soil Wetness
PI data	Wind Fetch
	FC2011

Together with the feature variables listed in Table 1, we added spectral vegetation indices calculated using a different combination of spectral bands as shown in Table 2.

Table 2. Spectral vegetative indices used as feature variables for classification.

Name	Equation	Notes	References
RVI	$\frac{\rho_{NIR}}{\rho_{RED}}$	Ratio Vegetation Index	Pearson & Miller (1972) [41]
NDVI	$\frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$	Normalised Difference Vegetative Index	Rouse, J.W., Jr. (1974) [42]
NDRE	$\frac{\rho_{NIR} - \rho_{RE}}{\rho_{NIR} + \rho_{RE}}$	Normalised Difference Red Edge Index	Gitelson et al. (1994) [43]
SAVI	$\frac{\rho_{NIR} + \rho_{RED} + L}{(1 + L)(\rho_{NIR} - \rho_{RED})}$	Soil Adjusted Vegetation Index	Huete, A.R. (1988) [44]
OSAVI	$\frac{1.5 * (\rho_{NIR} - \rho_{RED})}{\rho_{NIR} + \rho_{RED} + 0.16}$	Optimised Soil Adjusted Vegetation Index	Rondeaux et al. (1996) [45]
NLI	$\frac{\rho_{NIR}^2 - \rho_{RED}^2}{\rho_{NIR}^2 + \rho_{RED}^2}$	Non Linear Index	Goel & Qin (1994) [46]
MNLI	$\frac{(\rho_{NIR}^2 - \rho_{RED}^2) * (1 + L)}{\rho_{NIR}^2 + \rho_{RED}^2 + L}$	Modified Non Linear Index	Yang et al. (2008) [47]
BAI	$\frac{1}{(0.1 - \rho_{RED})^2 + (0.06 - \rho_{NIR})^2}$	Burn Area Index	Chuvieco et al. (2002) [48]

3.2.2. Spatial Relations

Researches have advocated the use of semivariograms for the improvement of object-based image analysis [32,49,50]. To explore spatial relations for GEOBIA, we applied semivariogram features.

A semivariogram is used to display the variability between data points as a function of distance. In remote sensing, semivariograms are calculated as half of the average squared difference between the reflectance value of a given spectral band separated by a given lag [51]. The semivariogram is computed as:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \quad (1)$$

where $\gamma(h)$ is the semivariance value at a certain lag distance h , and where $z(x_i)$ and $z(x_i + h)$ represent digital values at location x_i and $x_i + h$ respectively. $N(h)$ is the number of paired pixels at a lag distance h . The use of semivariogram features in GEOBIA follows the segmentation of the satellite image. The semivariogram features used for classification are presented in Table 3.

Table 3. Semivariogram features taken from [49].

Name	Equation	Notes
RVF	$\frac{\text{variance}}{\gamma_1}$	Ratio between total variance and first semivariance
RSF	$\frac{\gamma_2}{\gamma_1}$	Ratio between the first and the second semivariance
FDO	$\frac{\gamma_2 - \gamma_1}{h}$	First derivative near the origin
SDT	$\frac{\gamma_4 - 2\gamma_3 + \gamma_2}{h^2}$	Second derivative at third lag
FML	h_{\max_1}	First maximum lag value
MFM	$\gamma_{\max_1}^{\text{mean}} = \frac{1}{\max_1} \sum_{i=1}^{\max_1} \gamma_i$	Mean of the semivariogram values up to the first maximum
VFM	$\frac{1}{\max_1} \sum_{i=1}^{\max_1} (\gamma_i - \gamma_{\max_1}^{\text{mean}})^2$	Variance of the semivariogram values up to the first maximum
DMF	$\gamma_{\max_1}^{\text{mean}} - \gamma_1$	Difference between MFM and the first semivariance
RMM	$\frac{\gamma_{\max_1}}{\gamma_{\max_1}^{\text{mean}}}$	Ratio between the first local maximum semivariance and MFM
SDF	$\gamma_{\max_1} - 2\gamma_{\frac{\max_1}{2}} + \gamma_1$	Second-order difference between first lag and first maximum
AFM	$\frac{h}{2} \left(\gamma_1 + 2 \left(\sum_{i=2}^{\max_1-1} \gamma_i \right) + \gamma_{\max_1} \right) - (\gamma_1(h_{\max_1} - h_1))$	Second-order difference between first lag and first maximum
DMS	$h_{\max_2} - h_{\max_1}$	Distance between the first and the second local maxima
DMM	$h_{\min_1} - h_{\max_1}$	Distance between the first local maximum and the first local minimum
HA	$\frac{\frac{h}{2} (\gamma_{\max_1} + 2(\sum_{i=\max_1+1}^{\max_2-1} \gamma_{\max_2}))}{\frac{1}{2}(h_{\max_2} - h_{\max_1})(\gamma_{\max_2} + \gamma_{\max_1})}$	Hole area

3.2.3. Semantic Relations

In regular variograms, the variability between the observed numerical values of attributes are considered. In the case of semantic variograms, the difference between numerical values are replaced by semantic distances. The calculation of semantic distance is based on the semantic similarities between two classes. Thus, a semantic variogram is a measure of the variability between two classes based on the semantic similarities between classes at two different locations as opposed to their spatial distance [52,53]. The semantic variogram $\gamma_{SD}(h)$ for a lag distance h is computed as:

$$\gamma_{SD}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} sd[z(x_i); z(x_i + h)]^2 \quad (2)$$

where $N(h)$ is the number of pairs separated by h , and $sd[z(x); z(x+h)]$ is the semantic distance between the class of cell x_i and class of cell $x_i + h$.

3.2.4. Semantic Similarities

A semantic variogram is calculated based on the semantic distances between classes defined in an ontology [52]. Semantic distances can be quantified as the semantic similarity between two ontological classes based on the semantics associated between the objects [54,55]. Semantic similarity measures have been broadly categorised into hierarchy-based, information content based and feature-based [39].

- Hierarchy-based

The hierarchy-based similarity measure is a distance-based similarity measure that uses the conceptual hierarchy to calculate the distance between concepts. This distance is a count of the number of edges on the path or a count of the number of nodes in the path linking the two concepts. Thus it is also known as the path-based similarity or edge-counting similarity measures [39,55]. The semantic distance is measured by calculating the number of edges or the number of nodes that have to be traversed in a hierarchy from one concept to other. In Wu and Palmer's hierarchy-based measure [56], the similarity is calculated using the distance from the root to the common subsumer of C1 and C2 using the equation below.

$$Sim_{WP}(C1, C2) = \frac{2 \times len(root, C3)}{len(C1, C3) + len(C2, C3) + 2 \times len(root, C3)} \quad (3)$$

In Equation (3), Sim_{WP} is Wu and Palmer's [56] hierarchy-based similarity measure, C1 and C2 are the concepts whose semantic similarities are measured, C3 is the common subsumer of C1 and C2, root is the top concept in the hierarchy, and $len(root, C3)$ is the number of nodes on the path from concept C3 to the root concept.

- Information-content-based

Information content (IC) based similarity measures use a measure of how specific a concept is in a given ontology. If a concept is more specific, there will be high information content and inversely less information content with the more general concept. The ontology-based IC uses the ontology structure itself [57] which is defined in Equation (4) as below.

$$Sim_{IC}(C) = 1 - \frac{\log(num_{desc}(C)) + 1}{\log(max_{ont})} \quad (4)$$

where $Sim_{IC}(C)$ is the Similarity based on Information Content (IC), $num_{desc}(C)$ is the number of descendants for concept C and max_{ont} is the maximum number of concepts in the ontology.

- Feature-based

In an ontology, a class can be treated equivalent to another class if both classes have the same number of equivalent attributes. This means that the two classes are more highly similar when more common attributes exist between the classes. Thus the feature-based similarity measure is a degree of class similarity to another class. It is measured using the number of attributes that match between two classes. This approach consists of combining feature-based similarities within an ontology. The Tversky index is used to measure similarity based on the distinct features of class A to B, distinct features of class B to A, and common features of class A and B [58].

$$Sim_{Tversky}(C1, C2) = \frac{|A1 \cap A2|}{|A1 \cap A2| + \alpha |A1 \setminus A2| + \beta |A2 \setminus A1|} \quad (5)$$

where, $A1$ and $A2$ are the sets of attributes of classes $C1$ and $C2$; $|A1 \cap A2|$ is the total number of formal attributes shared by $C1$ and $C2$, $|A1|$ and $|A2|$ represent the number of formal attributes of $C1$ and $C2$; $\alpha = \beta = 1$ which is equivalent to Jaccard index.

4. Experiment: Case Study on Tasmanian Forests

4.1. Tasmanian Forest-Type Mapping

Accurate mapping of forest-type is a necessary step for forest inventory estimation, which in turn supports strategic forest management, carbon storage estimation, biological conservation and ecological restoration. Historically, Tasmania's forest-type mapping was carried out using stereoscopic interpretation of aerial photographs and referred to as photo-interpretation (PI) typing. PI-typing has served as a fundamental source of information for Tasmania's forest management [59]. Forest vegetation was segmented into patches that appeared visually homogeneous to highly skilled and experienced photo interpreters [59]. Each patch was assigned with a photo interpreted PI-type code that comprised a series of forest stand elements. These stand elements describe the forest associated with a patch using a standard set of characteristics such as species type, growth stage, structural group or forest group. The PI-type coding used in Tasmania is amenable to explicit formalisation and so provides an opportunity to investigate the modelling of domain knowledge into an ontology and the application of that ontology to multiple remote sensing data types to automate forest mapping. Table 4 shows the PI-typing associated with growth stage. Table 5 shows the PI-typing associated with structural grouping. The structural classification characterises forest stands into one of 12 broad categories according to their predominance of mature, regrowth, regenerated (regen) and non-eucalypt components (similar to STANDTYPE).

Table 4. Growth stage of Tasmanian forest.

Code	Name	Description
Y	Young Regeneration	Young native regeneration less than 20 years old.
R	Regrowth	Regrowth or regeneration older than 20 years.
M	Mature or Senescing	Mature or senescent (over-mature) forest.
U	Unknown	Unknown growth stage.
N	Not Applicable	Not applicable.

Table 5. Structural group of Tasmanian forest.

Code	Name/Description
MAT	Mature Eucalypt Forest, (with neither Regrowth nor aged eucalypt Regeneration)
MUR	Mature Eucalypt Forest with Unheighted Regrowth (and without aged eucalypt regeneration)
MAR	Mature Eucalypt with Aged Regeneration (from partial logging)
RGM	Unaged Regrowth Eucalypt with Mature (and without aged eucalypt regeneration)
REG	Pure Unaged Regrowth Eucalypt (and without mature or aged eucalypt regeneration)
RGA	Eucalypt Regrowth or older Aged Regeneration, with younger Aged Regeneration (from partial logging)
SIL	Even Aged Eucalypt Silvicultural Regeneration (An aged regeneration element, whether heighted or not, with no other mature or unaged eucalypt regrowth or aged eucalypt regeneration present)
UST	Unstocked Eucalypt Forest
RNF	Rainforest
ONF	Other Native Forest
PLN	Plantation
NOF	Non Forest

4.2. Assumptions

In this study, we have selected a structural group as our basis for classification. Among the structural group, we have selected Mature Eucalypt Forest (MAT), Pure Unaged Regrowth Eucalypt Forest (REG) and Even Aged Eucalypt Silvicultural Regeneration forest (SIL). The key reasons for this selection is that these forest types cover the majority of the study area and because of their high timber production value.

4.3. Ontology for Forest-Type Mapping

The forest-type mapping in this work is accomplished extending an Ontology-driven GEOBIA (O-GEOBIA) framework [5,6]. For this ontological framework, the ontology is developed using the structural group classification (Figure 2).

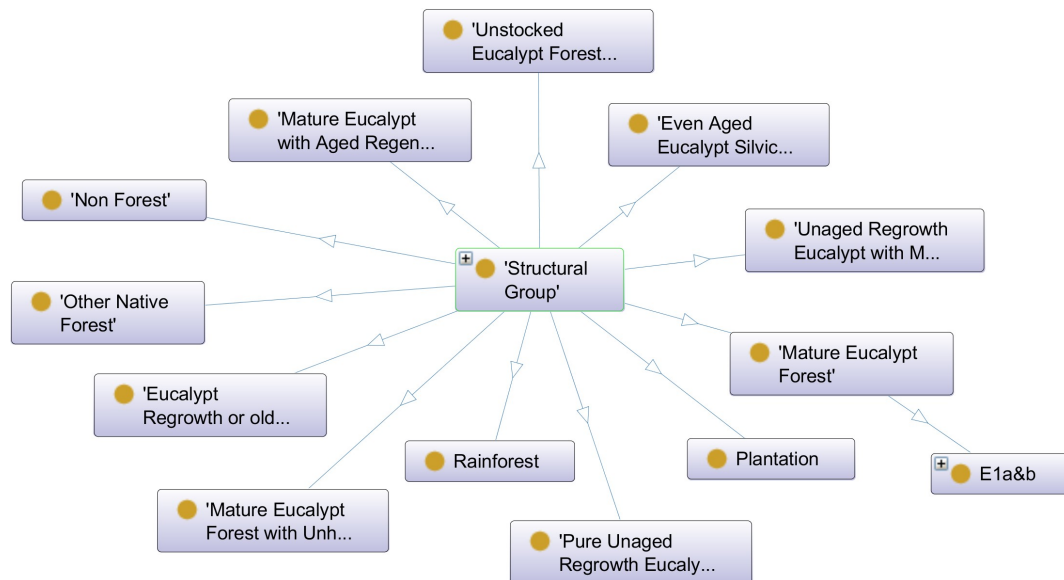


Figure 2. Structural group of a Tasmanian forest used to develop the ontology.

4.4. Study Area

The study area is located in northeast Tasmania, Australia, and is bounded between 517000E and 543000E and 5428000N and 5441500N and covers an area of approximately 356 km². A RapidEye satellite image of the study area is shown in Figure 3. The study area contains an almost complete representation of Tasmania's diverse forest types. The area has a complete coverage of Photo Interpretation (PI) and LiDAR data.

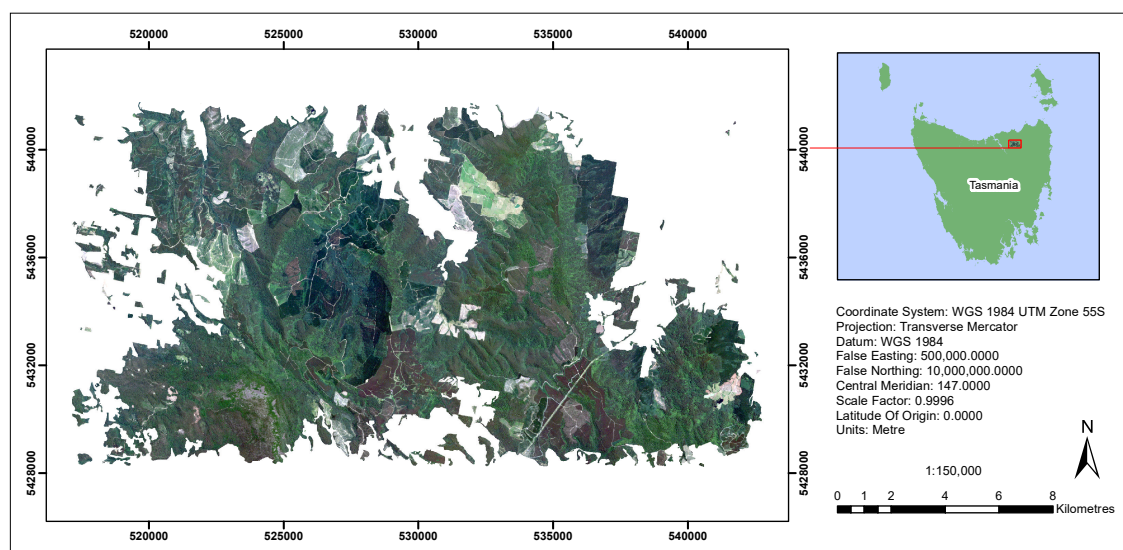


Figure 3. The location of the study area situated in the northeast of Tasmania, Australia.

4.5. Data

4.5.1. Satellite Image Data

A multispectral RapidEye dataset comprising 25 km × 25 km tiles (24 km + 500 m tile overlap) with UTM projection and WGS84 Datum was used. The ready to analyse imagery with radiometric, sensor and geometric corrections was acquired from RapidEye. The imagery has a spatial resolution of 5.0 m and includes five spectral bands (Table 6).

Table 6. RapidEye spectral band description.

Bands	Range
blue	(0.44–0.51 μm)
green	(0.52–0.59 μm)
red	(0.63–0.685 μm)
red-edge	(0.69–0.73 μm)
near-infra-red	(0.76–0.85 μm)

4.5.2. LiDAR Data

Airborne small footprint LiDAR data was acquired during January of 2010 and 2012 using an Optech Gemini discrete-return scanner operating at a 100 kHz laser repetition rate with a maximum scan angle off nadir of 15 degrees. The minimum pulse density was 200 per 10 square meters, and up to four returns were recorded per pulse. The laser scanner detects laser pulses reflected from the forest and terrain, providing information about the height and vertical stratification of the canopy elements. The intensity of each returned pulse also indicates the absorptive characteristics of the canopy elements, which may differ between different species. Both the height and intensity of pulse returns were used to create a number of variables. Vegetation height was derived by subtracting the highest returns from a digital surface model derived from the ground returns. A canopy surface height model and a surface intensity model were derived by fitting a b-spline curve to the highest and brightest vegetation return at 1 × 1 m spatial resolution. Percentiles 5–100% in 5% increments (e.g., ZPC90: 90 percentile of height value) and different moments such as Mean (e.g., CI_Mean: Mean value of Canopy Intensity), Standard Deviation (e.g., Z_SD: Standard Deviation value of height), Skew (e.g., Z_Skew: Skewness value of height), Kurtosis (Z_Kurt: Kurtosis value of height) and Range (e.g., Z_Range: Range value of height) were calculated for pulse height, pulse intensity, canopy surface height and canopy surface intensity. Additionally, the proportions of all signal returns and vegetation signal returns above certain heights were calculated for 1 m height increments from 1–5 m and 5 m height increments from 5–80 m. This produced 168 variables. Due to high levels of redundancy in this dataset, highly correlated LiDAR variables were removed with domain expert's recommendation, leaving 16 variables for inclusion in the models.

4.5.3. Photo Interpretation (PI) Data

For the past 50 years, PI-typing has served as a fundamental source of information for Tasmania's forest management [59]. PI-type codes provide a definition of height-class, crown density-class, stem-count class or condition-class that can be used to characterise each forest class. These forest classes have been grouped into the structural group presented in Table 5. The structural group MAT has 11 different forest classes categorised on the basis of their height, density and crown cover. For instance, "E1a&b" is one of the forest class where E = Mature Eucalypt; 1 = average height 55–76 m; a = 70–100% crown cover; b = 40–70% crown cover. Similarly, other forest structural groups are derived from the forest class defined in PI-data.

5. Implementation

The methodology described in Section 3 has been implemented for Tasmanian forest-type mapping using three different datasets. The overall steps are described in sub-sections below:

5.1. Data Fusion, Segmentation and Feature Extraction of Multi-Sensor Data

In this work, data fusion is carried out at the feature level. In feature level data fusion, at first the image is segmented into objects using segmentation techniques. Next, for each segmented image objects, features are extracted from different data sources. We fused three different types of data, namely a RapidEye satellite image (Tiff file), LiDAR data (RData file), and Photo Interpretation data (Shapefile) as shown in Figure 4.

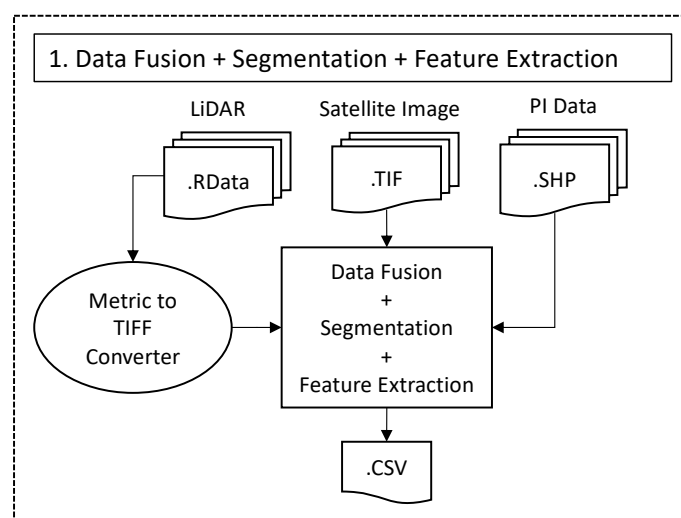


Figure 4. Data fusion, Segmentation and Feature extraction.

For segmentation and feature extraction, eCognition Developer Version 9.3.0 from Trimble, Germany was used. A chessboard segmentation technique was used for segmenting different forest-types in the RapidEye Satellite image. The object size parameter for chessboard segmentation was set to 6000 pixels (larger than image size) and thematic layer set to be taken from PI data. This ensures that the segmented image objects boundaries agree with the extent of PI data. From the RapidEye satellite image, different spectral indices were extracted as feature variables. LiDAR data was used to extract intensity, elevation, and their statistical metrics such as percentile and proportional values. For the calculation of semivariogram and related texture features, we used FETEX 2.0 from Geo-Environmental Cartography and Remote Sensing Research Group (CGAT), Spain [60]. PI-data was used to extract forest thematic features: class, structural group, growth stage and vegetation description.

5.2. Feature Selection

The fusion of multi-sensor data sources resulted in a large number of potential independent variables. With a high number of variables, the model will suffer from redundant features, overfitting and slow computation. However, there are two problems associated with reducing the dataset dimensionality: finding a minimal set of variables that are optimal for classification known as the 'minimal-optimal' problem and finding all variables relevant to the target variable known as the 'all-relevant' problem [61]. In this work, feature selection was performed using the Boruta package developed in R [62] (Figure 5). The Boruta algorithm [35] is implemented for finding all relevant variables.

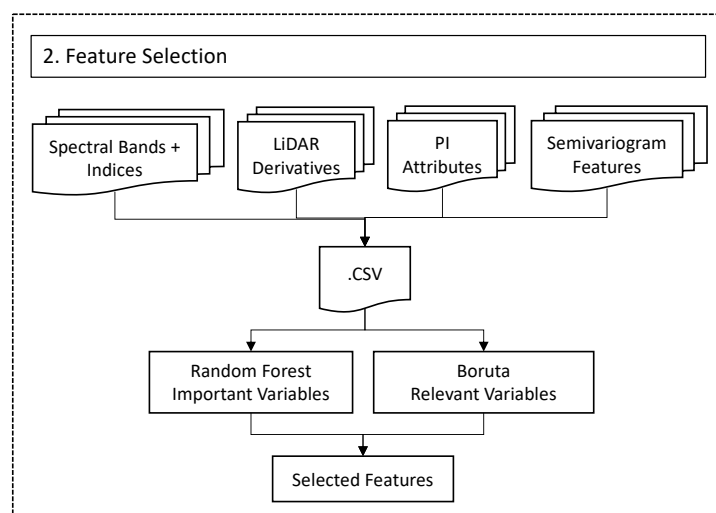


Figure 5. Feature Selection from multi-sensor data using machine learning techniques.

In the Boruta algorithm, duplicate copies of all the independent variables are created and shuffled. These duplicated variables are termed shadow variables. Next, a random forest classifier is used to identify the variable importance which results in a Z score. The Z score is the mean of accuracy loss divided by standard deviation of accuracy loss. The maximum Z score (MZSA) is calculated among those shadow variables. All the variables having importance lower than MZSA are tagged as unimportant and those higher than MZSA are tagged as important. The process is repeated until all the variables are tagged as important or unimportant. Based on the result from the Boruta algorithm, the important variables are treated as all relevant variables. To check the consistency and robustness of the model, we used k-fold cross validation approach to run Boruta. The dataset was randomly split into 10 equal size subsamples with 75% of data was used for training and 25% of data was used for validation. The selected all relevant variables serve as the input variables for ML techniques to extract classification rules.

5.3. Rules Generation

ML techniques are used to discover knowledge from data that are not conducive to human analysis, have high feature dimensions and a high number of predictor variables. Supervised data helps to identify potential classification rules. In our work, forest class information from PI-Data is used as supervised data to train the ML model. In a nutshell, we aim to leverage ML to automatically extract classification rules out of the available remote sensing datasets.

Random Forests (RF) as an ensemble algorithm can produce a very good predictive result but it acts as a black box model. With the thousands of decision trees as a forest, the ease of interpretation of a single model is lost. The inTrees framework uses the following steps to close the gap of model interpretability by converting the ensemble of models into a single model [63]. In this framework, the rules are extracted from each decision tree in the tree ensemble. The rules are then ranked, based on their frequency (measuring the rule's popularity), error (defining incorrectly classified instances), and length of the rule conditions (representing complexity). The rules are then pruned to remove irrelevant variable value pairs from the rule conditions. The selection of relevant and non-redundant conditions is performed using a feature selection approach. Finally, these processed rules are summarised into a simple set of if/then rules. In this work, we used the inTrees framework to prune large rulesets with redundant rules extracted from RF into simplified rules ready for the classification task (Figure 6). Such rules are transformed into SWRL (Semantic Web Rule Language) [64] to be used by the ontological reasoner for classification.

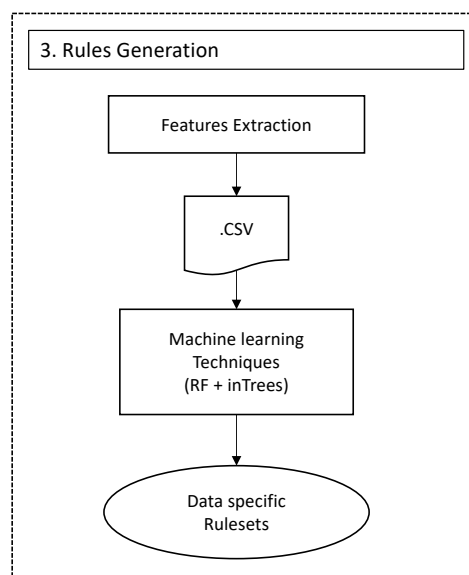


Figure 6. Use of Machine learning technique for classification rules generation.

5.4. Ontology-Based Image Classification

In the ontology-based image classification, we defined concepts, relations among them and instances to represent a domain of interest using the machine readable Web Ontology Language (OWL) [65]. The knowledge captured in a PI-type coding has not been organised in a formal machine-readable format to be used by forest planners in past applications [59]. In this work, we used the PI-type coding to model the Tasmanian forest domain knowledge for forest mapping. Subsequently, we extract potential instances as the segmented image objects from the data pre-processing module. The rules are defined in SWRL specification [64] as acquired from the rule generation module. The ontological framework [6] for the representation and reasoning over ontologies using Pellet reasoner [66] has been used, which executes the developed SWRL rules using the reasoning tools (Figure 7).

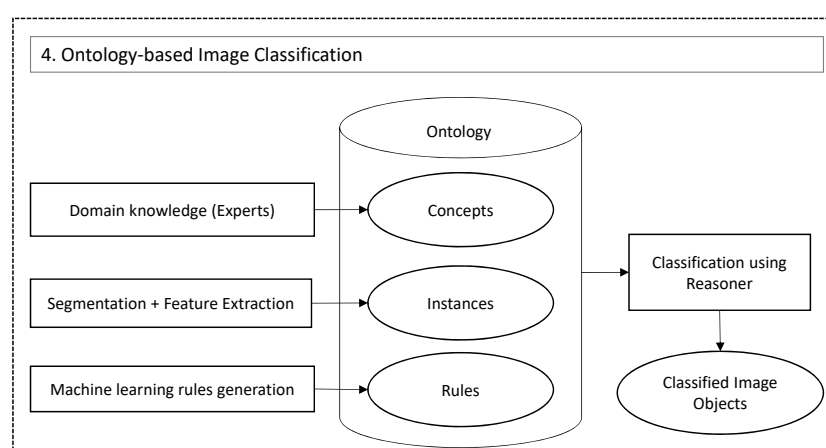


Figure 7. Ontology-based image classification framework.

6. Results

The classification results are based on the use of spectral indices, LiDAR derivatives and variogram features individually and finally with a combination of all three cases.

6.1. Feature Selection

The result from the application of the Boruta algorithm [35] to identify the relevant variables is presented in this section. The relevant feature selection experiment was implemented in five stages: (i) Using only spectral bands (Figure 8a); (ii) Using spectral band and vegetative indices based on the spectral bands (Figure 8b); (iii) Using LiDAR derivatives (Figure 8c); (iv) Using variogram features (Figure 8d); and (v) combining all the features from the previous four cases (Figure 9). In Figure 8a,c, all the variables have higher importance than shadow variables. Thus, in these cases all the variables are considered relevant and are represented by a green boxplot. The variables in the yellow boxplot shown in the Figures 8b,d and 9 are considered tentative variables whereas red boxplot variables are determined as unimportant attributes.

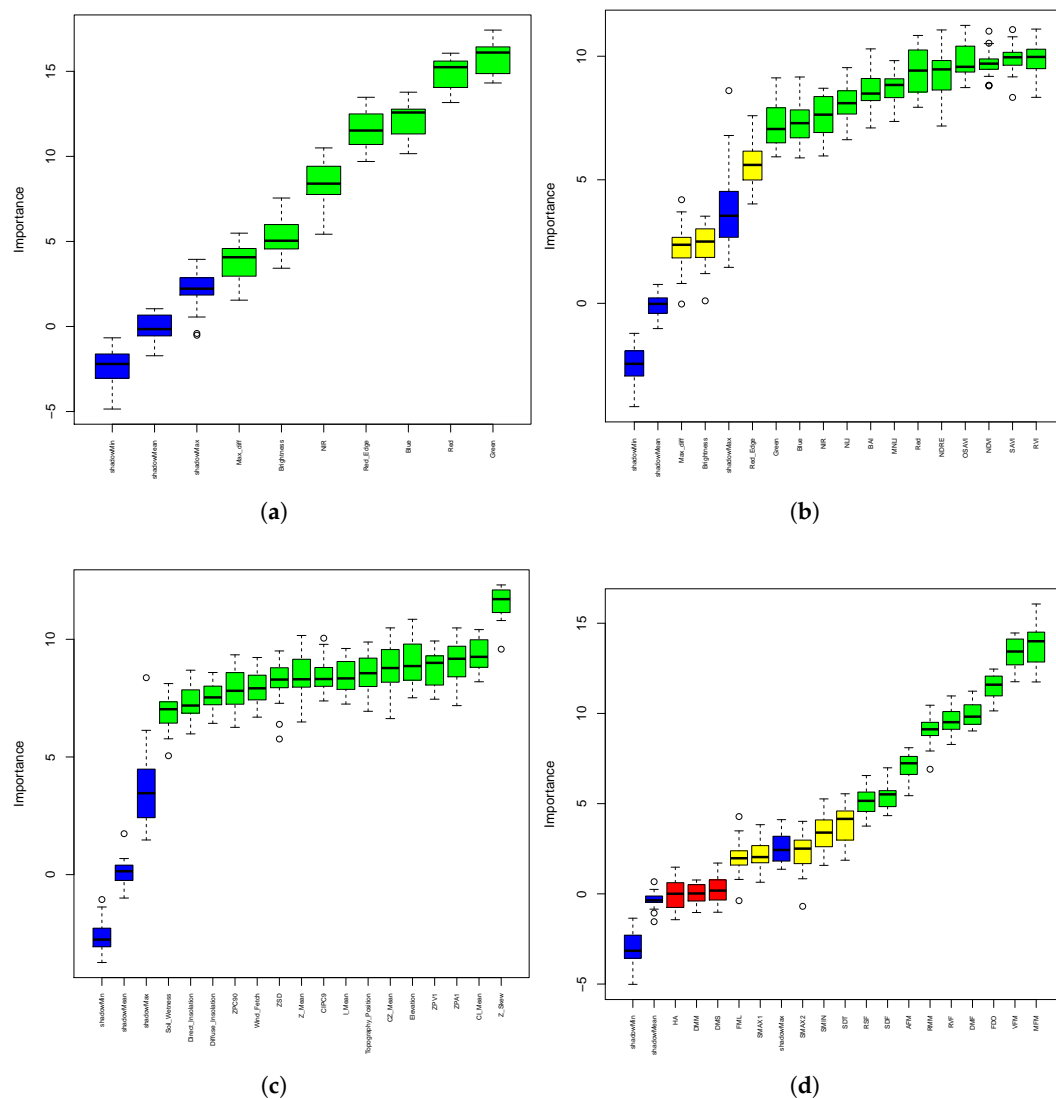


Figure 8. Boruta plot using (a) Spectral features (b) Spectral and Vegetative Indices (c) LiDAR derivatives (d) Variogram features. The x-axis shows the feature variables and y-axis shows importance in terms of Z-scores. In this figure, green boxplots are relevant variables, red boxplots are unimportant variables, yellow boxplots are tentative variables and the three blue boxplots represent maximum, median and minimum importance for shadow variables. Variables with an importance value lower than shadowMax are tagged as unimportant and higher than shadowMax are tagged as important.

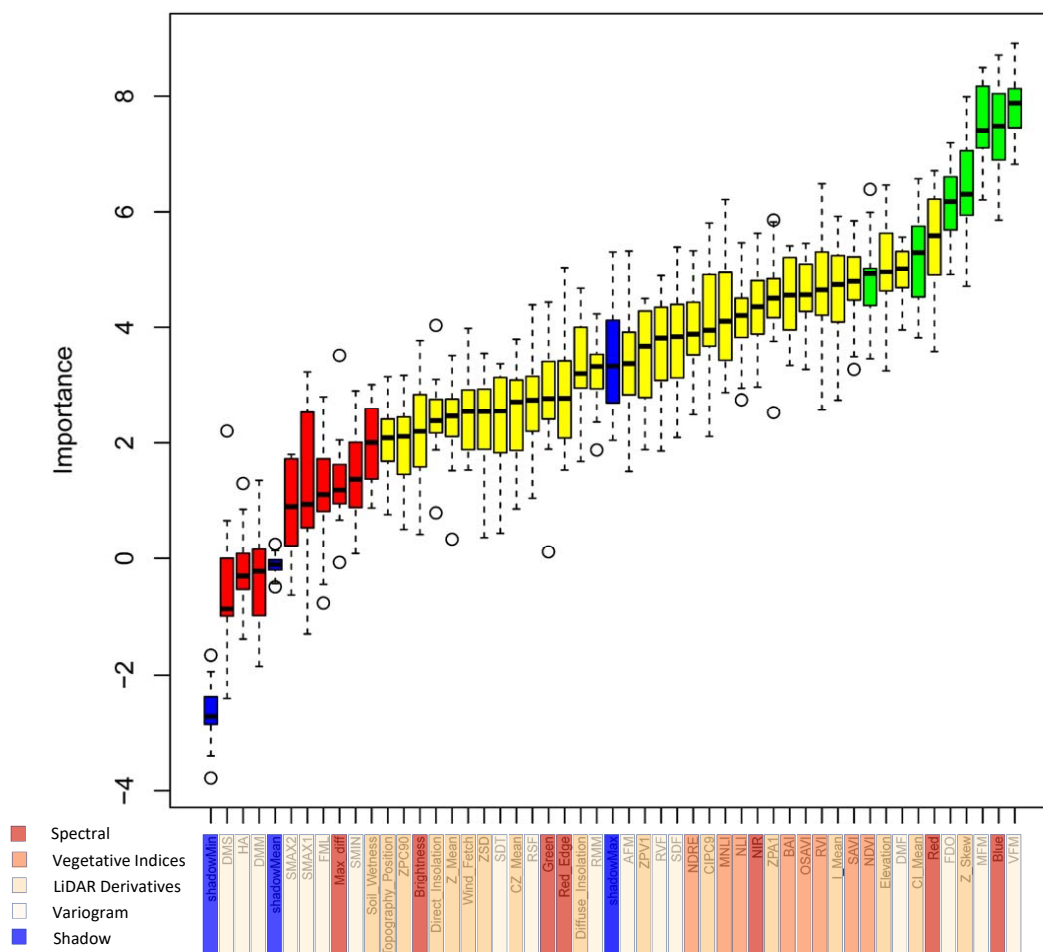


Figure 9. Boruta plot with all feature variables (spectral, vegetative indices, lidar and variogram). The x-axis shows the feature variables coloured according to data sources and the y-axis shows importance in terms of Z-scores. In this figure, green boxplots are relevant variables, red boxplots are unimportant variables, yellow boxplots are tentative variables and the three blue boxplots represent maximum, median and minimum importance for shadow variables. Variables with an importance value lower than shadowMax are tagged as unimportant and higher than shadowMax are tagged as important.

Figure 9 shows that 7 variables (Blue, CI_Mean, FDO, MFM, NDVI, Z_Skew, VFM) are confirmed important, 9 variables are confirmed unimportant and remaining 32 as tentative variables. The number of classifier (Random Forest) runs during the Boruta algorithm execution is limited by the maxRuns argument (maxRuns = 18). This leaves attributes that need to be judged important or unimportant are marked as tentative variables. A diagnostic plot depicting the fluctuation of variable importance after several iterative runs of the Boruta algorithm is shown in Figure 10a. In Figure 10b, a scatter plot shows the importance of each variables at each classifier run. In the plot, the green lines with higher importance than shadowMax variables represent relevant variables. In the first few runs some important variables (CI_Mean, NDVI) are below shadowMax variable as shown in Figure 10c and an unimportant variable (SMIN) is above shadowMax variable as shown in Figure 10d. Thus, the Boruta algorithm runs multiple Random Forest before arriving at a statistically significant decision. The selection criteria for the maxRun parameter is the Random Forest classifier run that results in the minimum number of variables and maximum classification accuracy. For instance, maxRun limits of 18 and 500 resulted in 7 and 26 confirmed important variables but with the same classification accuracy of 82.57%.

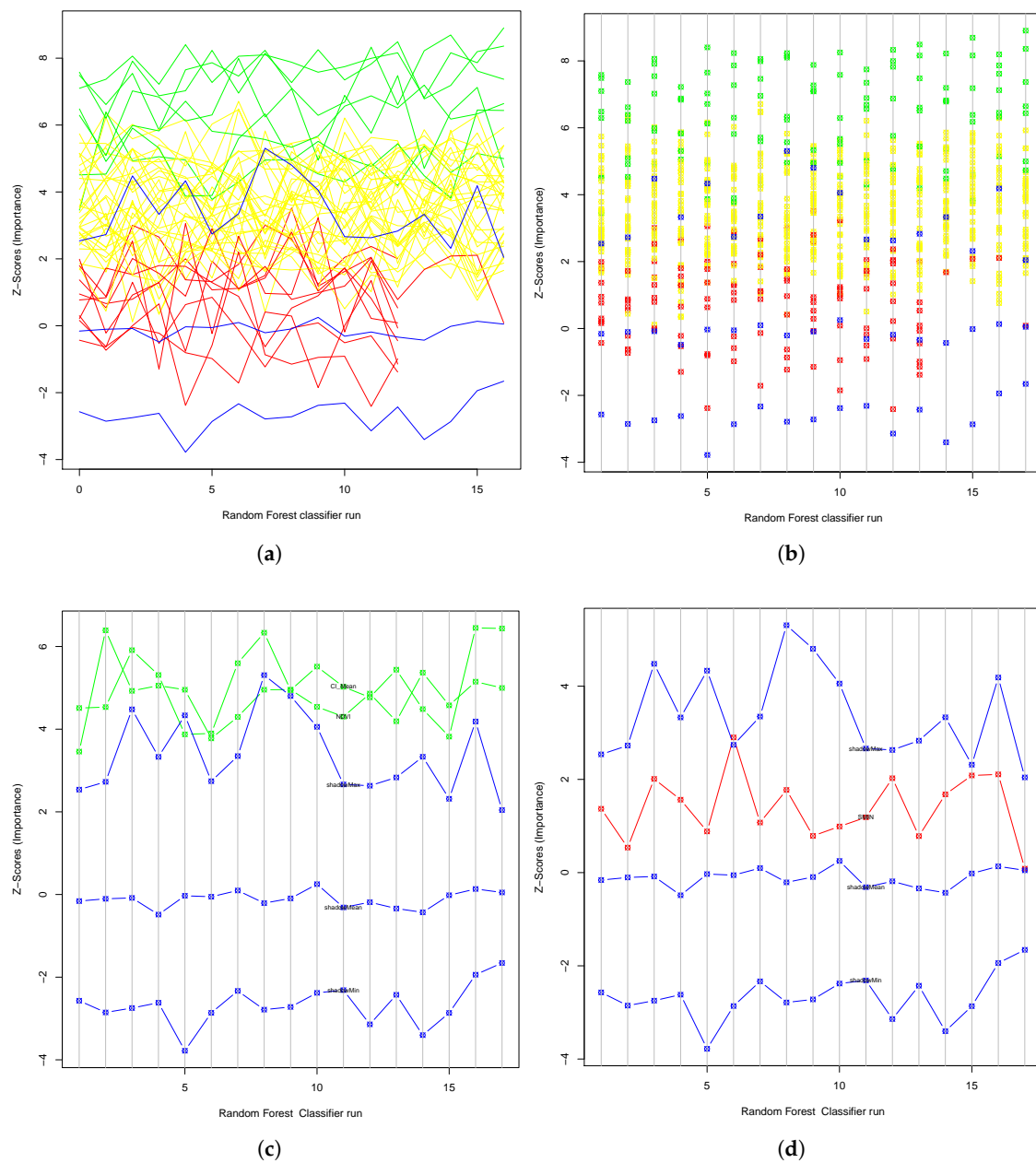


Figure 10. Diagnostic plot of the Boruta algorithm showing (a) line plot of Z-scores (importance) at different Random Forest run for each variable (b) scatter plot of Z-scores (importance) at different Random Forest run for each variable (c) important variables appearing below shadowMax at certain instance of Random Forest run (d) unimportant variables appearing above shadowMax at certain instance of Random Forest run. In the figures, green lines and points represent important variables, red lines and points represent unimportant variables, yellow lines and points represent tentative variables and the three blue plots represent maximum, median and minimum importance for shadow variables.

6.2. Classification Accuracy Assessment

A confusion matrix is used to assess image classification accuracy. The matrix is created for three different forest-type classes MAT, REG and SIL where the ground truth data is taken from PI data. The accuracy assessment is carried out as 5 experiments based on individual features and combination of all, as shown in Table 7.

Table 7. Confusion matrix between Spectral, Vegetative indices, LiDAR, and Variogram. In the table, classes M, R and S represents MAT, REG and SIL respectively. The OA represents Overall Accuracy.

Class	Spectral			Spectral + Indices			LiDAR			Variogram			All		
	M	R	S	M	R	S	M	R	S	M	R	S	M	R	S
M (MAT)	3	4	0	4	3	0	6	9	0	2	3	0	6	2	0
R (REG)	20	81	1	19	82	1	17	74	1	21	82	1	17	83	1
S (SIL)	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
OA	77.06%			78.90%			73.39%			77.06%			81.65%		

We experimented to compare the overall classification accuracy based on the features from different sensors. The overall accuracy of spectral variables extracted from satellite imagery is 77.06%. Next, we calculated standard vegetative indices using different spectral bands, which improved the classification accuracy to 78.90%. Using LiDAR data without spectral features achieved an overall accuracy to 73.39%. In the variogram based experiment, the accuracy is increased to 77.06%, similar to that with spectral feature based first experiment. In the final experiment, spectral, LiDAR and variogram features are used which showed the highest accuracy of 81.65%. The results suggest that our aim to use data fusion to increase the number of feature variables for higher classification accuracy is achieved. To tackle the feature dimension issue, the Boruta algorithm was then applied to extract relevant variables. The subsequent classification, carried out using relevant variables, resulted in a slightly higher accuracy of 82.57% (Table 8). Only one SIL plot is available in the given test dataset which was not correctly classified in any of the experiments. However, we included this in our experimental design as SIL is one of the representative forest types in Tasmania.

Table 8. Confusion matrix for classification result based on all the variables available and relevant ones extracted from the Boruta algorithm.

Class	All Available Variables			Relevant Variables		
	M	R	S	M	R	S
M (MAT)	6	2	0	8	3	0
R (REG)	17	83	1	15	82	1
S (SIL)	0	0	0	0	0	0
OA	81.65%			82.57%		

6.3. Semantic Similarity Assessment

The first step to determine the semantic similarity between classes is to find out the common attributes. Figure 11 shows the sharing of attributes for each class.

The hierarchy-based similarities investigated in this work are based on Wu and Palmer [56] using Equation (3). For the feature-based similarities calculation, the Tversky index [58] using Equation (5) is used.

Our results indicated that feature-based similarity measures were more capable of differentiating among classes than hierarchy-based methods (Table 9). The Wu and Palmer's hierarchy-based similarity has the same index value of 0.25 for all the pairs of classes without being able to detect any dissimilarity. This is explained by the equal hop distance between classes in the hierarchy. The Tversky feature-based similarity measure showed significant differences between classes. Classes pairs that match a higher number of attributes result in a higher similarity index value. The results show that classes MAT and REG have a higher index with more attributes matched.

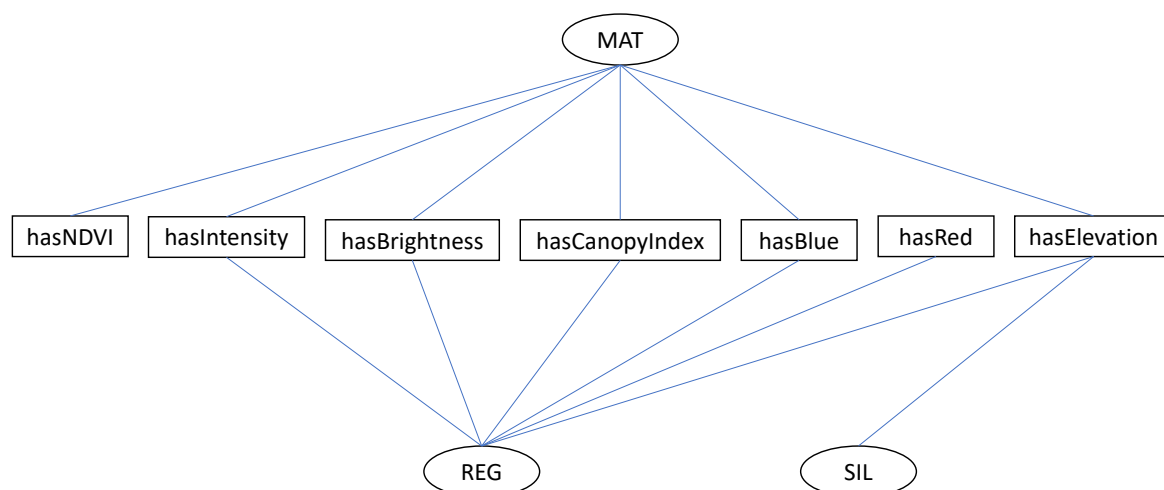


Figure 11. Ontological graph showing the concepts and associated attributes. This graph shows the sharing of the attributes of different concepts. The concept “MAT” and “REG” has the following common attributes: {hasCanopyIntensity; hasIntensity; hasBrightness; hasBlue}. The attributes “hasNDVI” and “hasRed” only belong to concepts “MAT” and “REG” respectively. The attributes “hasElevation” belongs to all 3 concepts.

Table 9. Comparison between similarity measures.

		Wu & Palmer	Tversky
MAT	REG	0.25	0.72
	SIL	0.25	0.17
REG	MAT	0.25	0.72
	SIL	0.25	0.17
SIL	MAT	0.25	0.17
	REG	0.25	0.17

7. Discussion

7.1. Importance of Feature Selection in the Fused Multi-Sensor Data

With the fusion of multi-sensor data, the feature dimension increases and provides more variables available to use in classification. The fusion process introduces non-relevant and redundant variables that increases complexity and computational load. In tackling such circumstances, a feature selection algorithm is used to reduce the number of variables without compromising overall classification accuracies. In Table 8, we show how the classification accuracy is increased by 0.92% even when the feature variables are reduced from 48 to 7.

Boruta offers an improvement over the Random Forests variable importance measure. In Random Forests, the calculated Z score is not directly related to the statistical significance of the variable importance. Boruta runs Random Forest on both original and random attributes and computes the importance of all variables. Since the whole process is dependent on permuted copies, we repeat the random permutation procedure to get statistically robust results for our fused datasets. The result presented in Table 8 shows how the classification accuracy is increased from 81.65% to 82.57% when using simple RF with all features over relevant features extracted using Boruta. Considering the scope of this work, no comparative evaluation of Boruta [35] with other feature selection mechanism such as Altmann [67], r2VIM (Recurrent relative variable importance) [68] or Vita [69] was carried out. The current research shows that Boruta efficiently identifies relevant variables in high-dimensional datasets [70,71].

7.2. Evaluation of Semivariogram Features

The semivariogram has been applied in remote sensing to extract texture features and the spatial structure for image classification. The usage of semivariograms varies from different types of sensor data to different applications such as forest structure mapping [72] or classification of land use [49], land cover [73] or vegetation communities [74]. With the advent of GEOBIA, semivariograms have also been implemented in object-based image analysis [32,49,50]. To achieve harmony with GEOBIA, the extraction of the semivariogram is performed for image segments instead for a certain size window or kernel. Thus, we can claim that this is an object-based semivariogram as the calculation of semivariogram is done within the boundary of each image segment. Within the extent of a segmented object, a sequence of semivariance values is calculated, from which variogram variables will be extracted. However, we have not experimented the variation of segments area and robustness in resulting scenarios while selecting variogram variables. This is not tested statistically in this study considering it is beyond the scope of the study.

In this study, we tested the efficiency of semivariogram derived features as proposed by [49]. The classification accuracy of object-based image classification is compared between using 'semivariogram features' vs. 'other sets of features extracted from spectral and lidar data'. The result shows that the overall accuracy when using only the semivariogram derived features is 77.06%, which is equivalent to that of spectral features (77.06%). The feature selection algorithm Boruta showed the slope of the first two lags (FDO), mean and variance of the semivariogram values up to the first maximum (MFM and VFM) to be relevant variables.

7.3. Selection of Semantic Similarities for Multi-Sensor Remote Sensing Data

GEOBIA is intended to align with the methods by which humans identify and classify objects [2,75,76]. For the success of an ontological GEOBIA framework, the ontology needs to be developed with a focus on human activities in geographical space [23]. In image interpretation, there is the lack of assessment of semantic likeness that occurs between image object classes. Ontology can measure similarity that is based on semantics [77]. Ontology-based semantic similarity quantifies how taxonomically two classes based on their features are similar. In this regard, applications developed based on ontological domain knowledge require quantification of relationships between ontological concepts [54].

Nevertheless, there exist different ontology-based semantic similarity approaches. To understand and select the suitable method for a specific application is a challenge. For determining the method that suits our forestry mapping application, different semantic similarity measures were studied and tested. This work develops an innovation purely in ontological space—in calculating a semantic similarity measure. In ontology-based semantic similarity measures, there are edge-counting, features based and information content methods. The computation of these methods is simple and efficient as they only exploit the semantic network provided by the ontology.

Among these, the edge-counting similarity measure is the simplest and is computationally efficient [39]. However its similarity index is not suitable for the ontological model with the simple hierarchical structure as it cannot exploit complex semantics hidden within the class. This was clearly shown in the result presented in the Table 9 where the semantic similarity measures calculated were the same for all the classes in the same hierarchy.

The Information Content approach is based on capturing implicit semantic information as a function of concept distribution in corpora [39]. Such an approach is useful in natural language processing work where the association between the words found in a corpus and concepts are used to compute accurate concept appearance frequencies. In our work, where the image classification is carried out on the basis of feature attributes, the information content approach is not applicable.

Feature-based methods try to overcome the limitations of hierarchy-based measures by considering ontological features of each class. Feature-based approaches thus rely on a taxonomic hierarchy, relationships and attributes to determine the similarity between classes. Table 9 shows

how a feature-based approach (Tversky) makes the distinction between different classes compared to edge-counting (Wu & Palmer). Similar results were presented where feature-based semantic measures performed better than the edge-counting measures [77].

7.4. Limitations

In this work, we employed semantic similarities based on the ontological data. The applicability and accuracy of the similarity measures depend on the availability of well-defined domain ontologies. This means that poor construction of the domain ontology will result in non-robust semantic similarities between the domain classes. Also, the similarity is calculated based on the taxonomical hierarchy relations. The non-taxonomic relations (e.g., object x is part of object y) that can help to determine better similarity measures are missing. The discovery of non-taxonomic relations is a fundamental point in domain knowledge construction and with its addition, semantic similarity measure will be improved [78].

Among the different semantic similarity approach, we used a feature based semantic similarity approach. Each feature used in finding the similarities can have a varied contribution in classifying different classes. This phenomenon of feature contribution per class has not been considered in this work. To overcome this limitation, each feature can be given a certain weight based on the contribution; this is a topic for future research.

8. Conclusions

This research has extended an ontology based GEOBIA framework described in [6] to the case of a data fusion environment. The innovation in this study is that multi-sensor data has been fused into an integrated ontological image analysis framework. The developed framework incorporates spectral, spatial, textural and semantic features. The issue of high feature dimensionality raised with data fusion is addressed using a machine learning technique, in our case the Boruta algorithm. The algorithm determines the relevant features used for classification. Semantic similarity techniques are exploited for the characterisation of different forest-type classes. A semantic variogram is used to show the spatial and semantic relations of the different forest-type classes. The GEOBIA community and the science of O-GEOBIA can benefit from these types of extension within a GEOBIA methodology to tackle the issues of multi-sensor data fusion.

Author Contributions: S.R. and J.A. conceived the project; S.R. processed the data and wrote the first draft of the manuscript; J.A., J.O., A.L. and R.M. commented on the manuscript and supervised the project.

Funding: This research received no external funding.

Acknowledgments: The authors would like to acknowledge the University of Tasmania, Discipline of Geography and Spatial Sciences for logistics support and Sustainable Timber Tasmania for providing datasets and fruitful discussions. Sachit Rajbhandari is supported by an Australian Government Research Training Program Scholarship. Authors would like to thank Ola Ahlqvist of The Ohio State University and Ashton Shortridge of Michigan State University for sharing their published work and sample codes on Semantic Variograms.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DEM	Digital Elevation Model
DSM	Digital Surface Model
FC2011	Forest Class 2011
GEOBIA	Geographic Object-Based Image Analysis
IC	Information Content
inTrees	interpretable Trees
LAS file	LASer file
LiDAR	Light Detection and Ranging
MAT	Mature Eucalypt Forest

ML	Machine Learning
MZSA	Maximum Z Score
O-GEOBIA	Ontology-driven Geographic Object-Based Image Analysis
OWL	Web Ontology Language
PI	Photo Interpretation
REG	Pure Unaged Regrowth Eucalypt Forest
r2VIM	Recurrent relative variable importance
RF	Random Forests
SIL	Even Aged Eucalypt Silvicultural Regeneration forest
SWRL	Semantic Web Rule Language
UTM	Universal Transverse Mercator
WGS84	World Geodetic System 1984

References

1. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Queiroz Feitosa, R.; van der Meer, F.; van der Werff, H.; van Coillie, F.; et al. Geographic Object-Based Image Analysis - Towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)] [[PubMed](#)]
2. Addink, E.A.; Van Coillie, F.M.B.; De Jong, S.M. Introduction to the GEOBIA 2010 special issue: From pixels to geographic objects in remote sensing image analysis. *Int. J. Appl. Earth Observ. Geoinf.* **2012**, *15*, 1–6. [[CrossRef](#)]
3. Argyridis, A.; Argialas, D.P. Building change detection through multi-scale GEOBIA approach by integrating deep belief networks with fuzzy ontologies. *Int. J. Image Data Fusion* **2016**, *7*, 148–171. [[CrossRef](#)]
4. Arvor, D.; Durieux, L.; Andrés, S.; Laporte, M.A. Advances in Geographic Object-Based Image Analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 125–137. [[CrossRef](#)]
5. Rajbhandari, S.; Aryal, J.; Osborn, J.; Lucieer, A.; Musk, R. Employing Ontology to Capture Expert Intelligence within GEOBIA: Automation of the Interpretation Process. In *Remote Sensing and Cognition: Human Factors in Image Interpretation*; White, R., Coltekin, A., Hoffman, R., Eds.; Book Section 8; CRC Press: Boca Raton, FL, USA, 2018; pp. 151–170.
6. Rajbhandari, S.; Aryal, J.; Osborn, J.; Musk, R.; Lucieer, A. Benchmarking the Applicability of Ontology in Geographic Object-Based Image Analysis. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 386. [[CrossRef](#)]
7. Schmitt, M.; Zhu, X.X. Data Fusion and Remote Sensing: An ever-growing relationship. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 6–23. [[CrossRef](#)]
8. Dong, J.; Zhuang, D.; Huang, Y.; Fu, J. Advances in Multi-Sensor Data Fusion: Algorithms and Applications. *Sensors* **2009**, *9*, 7771–7784. [[CrossRef](#)] [[PubMed](#)]
9. Lu, M.; Chen, B.; Liao, X.; Yue, T.; Yue, H.; Ren, S.; Li, X.; Nie, Z.; Xu, B. Forest Types Classification Based on Multi-Source Data Fusion. *Remote Sens.* **2017**, *9*, 1153. [[CrossRef](#)]
10. Sadjadi, F. Comparative Image Fusion Analaysais. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)—Workshops, San Diego, CA, USA, 21–23 September 2005; p. 8.
11. Zhang, J. Multi-source remote sensing data fusion: Status and trends. *Int. J. Image Data Fusion* **2010**, *1*, 5–24. [[CrossRef](#)]
12. Johansen, K.; Tiede, D.; Blaschke, T.; Arroyo, L.A.; Phinn, S. Automatic Geographic Object Based Mapping of Streambed and Riparian Zone Extent from LiDAR Data in a Temperate Rural Urban Environment, Australia. *Remote Sens.* **2011**, *3*, 1139–1156.
13. Kempeneers, P.; Sedano, F.; Seebach, L.; Strobl, P.; San-Miguel-Ayanz, J. Data Fusion of Different Spatial Resolution Remote Sensing Images Applied to Forest-Type Mapping. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4977–4986. [[CrossRef](#)]
14. Tönjes, R.; Growe, S.; Bückner, J.; Liedtke, C.E. Knowledge-based interpretation of remote sensing images using semantic nets. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 811–821.
15. Durand, N.; Derivaux, S.; Forestier, G.; Wemmert, C.; Gancarski, P.; Boussaid, O.; Puissant, A. Ontology-based object recognition for remote sensing image interpretation. In Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence, Patras, Greece, 29–31 October 2007; pp. 472–479.

16. Costa, G.; Feitosa, R.; Fonseca, L.; Oliveira, D.; Ferreira, R.; Castejon, E. Knowledge-based interpretation of remote sensing data with the InterIMAGE system: Major characteristics and recent developments. In Proceedings of the 3rd GEOBIA, Ghent, Belgium, 29 June–2 July 2010.
17. Mundy, J.L.; Dong, Y.; Gilliam, A.; Wagner, R. The Semantic Web and Computer Vision: Old AI Meets New AI. In Proceedings of the Automatic Target Recognition XXVIII, Orlando, FL, USA, 30 April 2018; Volume 10648, p. 8.
18. Belgiu, M.; Hofer, B.; Hofmann, P. Coupling formalized knowledge bases with object-based image analysis. *Remote Sens. Lett.* **2014**, *5*, 530–538. [[CrossRef](#)]
19. Gu, H.; Li, H.; Yan, L.; Liu, Z.; Blaschke, T.; Soergel, U. An Object-Based Semantic Classification Method for High Resolution Remote Sensing Imagery Using Ontology. *Remote Sens.* **2017**, *9*, 329. [[CrossRef](#)]
20. Bittner, T.; Winter, S. *On Ontology in Image Analysis*; Integrated Spatial Databases; Springer: Berlin/Heidelberg, Germany, 2000; pp. 168–191.
21. Frank, A.U. Tiers of ontology and consistency constraints in geographical information systems. *Int. J. Geogr. Inf. Sci.* **2001**, *15*, 667–678. [[CrossRef](#)]
22. Winter, S. Ontology: Buzzword or paradigm shift in GI science? *Int. J. Geogr. Inf. Sci.* **2001**, *15*, 587–590. [[CrossRef](#)]
23. Kuhn, W. Ontologies in support of activities in geographical space. *Int. J. Geogr. Inf. Sci.* **2001**, *15*, 613–631. [[CrossRef](#)]
24. Agarwal, P. Ontological considerations in GIScience. *Int. J. Geogr. Inf. Sci.* **2005**, *19*, 501–536. [[CrossRef](#)]
25. Blaschke, T.; Lang, S.; Lorup, E.; Strobl, J.; Zeil, P. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. *Environ. Inf. Plan. Politics Public* **2000**, *2*, 555–570.
26. Mezaris, V.; Kompatsiaris, I.; Strintzis, M.G. An ontology approach to object-based image retrieval. In Proceedings of the 2003 International Conference on Image Processing (Cat. No.03CH37429), Barcelona, Spain, 14–17 September 2003; Volume 2, p. II-511.
27. Hay, G.J.; Castilla, G. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*; Blaschke, T., Lang, S., Hay, G.J., Eds.; Book Section Chapter 4; Lecture Notes in Geoinformation and Cartography; Springer: Berlin/Heidelberg, Germany, 2008; pp. 75–89.
28. Andrés, S.; Pierkot, C.; Arvor, D. Towards a Semantic Interpretation of Satellite Images by Using Spatial Relations Defined in Geographic Standards. In Proceedings of the Fifth International Conference on Advanced Geographic Information Systems, Applications, and Services, Nice, France, 24 February–1 March 2013.
29. Gruber, T.R. A translation approach to portable ontology specifications. *Knowl. Acquis.* **1993**, *5*, 199–220. [[CrossRef](#)]
30. Pohl, C.; Van Genderen, J.L. Review article Multisensor image fusion in remote sensing: Concepts, methods and applications. *Int. J. Remote Sens.* **1998**, *19*, 823–854. [[CrossRef](#)]
31. Huete, A.R. Vegetation indices, remote sensing and forest monitoring. *Geogr. Compass* **2012**, *6*, 513–532. [[CrossRef](#)]
32. Wu, X.; Peng, J.; Shan, J.; Cui, W. Evaluation of semivariogram features for object-based image classification. *Geo-Spatial Inf. Sci.* **2015**, *18*, 159–170. [[CrossRef](#)]
33. Kohavi, R.; John, G.H. Wrappers for feature subset selection. *Artif. Intell.* **1997**, *97*, 273–324. [[CrossRef](#)]
34. Blum, A.L.; Langley, P. Selection of relevant features and examples in machine learning. *Artif. Intell.* **1997**, *97*, 245–271. [[CrossRef](#)]
35. Kursu, M.B.; Rudnicki, W.R. Feature selection with the Boruta package. *J. Stat. Softw.* **2010**. [[CrossRef](#)]
36. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
37. Langley, P.; Simon, H.A. Applications of machine learning and rule induction. *Commun. ACM* **1995**, *38*, 54–64. [[CrossRef](#)]
38. Ben-David, A.; Mandel, J. Classification Accuracy: Machine Learning vs. Explicit Knowledge Acquisition. *Mach. Learn.* **1995**, *18*, 109–114. [[CrossRef](#)]
39. Sánchez, D.; Batet, M.; Isern, D.; Valls, A. Ontology-based semantic similarity: A new feature-based approach. *Expert Syst. Appl.* **2012**, *39*, 7718–7728. [[CrossRef](#)]

40. Cross, V.; Xueheng, H. Fuzzy set and semantic similarity in ontology alignment. In Proceedings of the 2012 IEEE International Conference on Fuzzy Systems, Brisbane, QLD, Australia, 10–15 June 2012; pp. 1–8.
41. Pearson, R.L.; Miller, L.D. Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie. In Proceedings of the Eighth International Symposium on Remote Sensing of Environment, Ann Arbor, MI, USA, 2–6 October 1972; p. 1355.
42. Rouse, J.W., Jr. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; FAO: Rome, Italy, 1973.
43. Gitelson, A.; Merzlyak, M.N. Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *J. Plant Physiol.* **1994**, *143*, 286–292. [\[CrossRef\]](#)
44. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [\[CrossRef\]](#)
45. Rondeaux, G.; Steven, M.; Frederic, B. Optimization of Soil-Adjusted Vegetation Indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [\[CrossRef\]](#)
46. Goel, N.S.; Qin, W. Influences of canopy architecture on relationships between various vegetation indices and LAI and FPAR: A computer simulation. *Remote Sens. Rev.* **1994**, *10*, 309–347. [\[CrossRef\]](#)
47. Yang, Z.J.; Tsubakihara, H.; Kanae, S.; Wada, K.; Su, C.Y. A novel robust nonlinear motion controller with disturbance observer. *IEEE Trans. Control Syst. Technol.* **2008**, *16*, 137–147. [\[CrossRef\]](#)
48. Chuvieco, E.; Martín, M.P.; Palacios, A. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *Int. J. Remote Sens.* **2002**, *23*, 5103–5110. [\[CrossRef\]](#)
49. Balaguer, A.; Ruiz, L.A.; Hermosilla, T.; Recio, J.A. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. *Comput. Geosci.* **2010**, *36*, 231–240. [\[CrossRef\]](#)
50. Powers, R.P.; Hermosilla, T.; Coops, N.C.; Chen, G. Remote sensing and object-based techniques for mapping fine-scale industrial disturbances. *Int. J. Appl. Earth Observ. Geoinf.* **2015**, *34*, 51–57. [\[CrossRef\]](#)
51. Atkinson, P.M.; Lewis, P. Geostatistical classification for remote sensing: An introduction. *Comput. Geosci.* **2000**, *26*, 361–371. [\[CrossRef\]](#)
52. Ahlqvist, O.; Shortridge, A. Spatial and semantic dimensions of landscape heterogeneity. *Landsc. Ecol.* **2010**, *25*, 573–590. [\[CrossRef\]](#)
53. Ahlqvist, O.; Shortridge, A. Characterizing Land Cover Structure with Semantic Variograms. In *Progress in Spatial Data Handling: 12th International Symposium on Spatial Data Handling*; Riedl, A., Kainz, W., Elmes, G.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 401–415.
54. Gan, M.; Dou, X.; Jiang, R. From Ontology to Semantic Similarity: Calculation of Ontology-Based Semantic Similarity. *Sci. World J.* **2013**, *2013*, 11. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Cross, V.; Yu, X.; Hu, X. Unifying ontological similarity measures: A theoretical and empirical investigation. *Int. J. Approx. Reason.* **2013**, *54*, 861–875. [\[CrossRef\]](#)
56. Wu, Z.; Palmer, M. Verbs Semantics and Lexical Selection. In Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics (ACL '94), Stroudsburg, PA, USA, 27–30 June 1994; pp. 133–138.
57. Seco, N.; Veale, T.; Hayes, J. An intrinsic information content metric for semantic similarity in WordNet. In Proceedings of the 16th European Conference on Artificial Intelligence, Valencia, Spain, 22–27 August 2004; pp. 1089–1090.
58. Tversky, A. Features of similarity. *Psychol. Rev.* **1977**, *84*, 327–352. [\[CrossRef\]](#)
59. Stone, M.G. Forest-type mapping by photo-interpretation: A multi-purpose base for Tasmania's forest management. *Tasforests* **1998**, *10*, 1–15.
60. Ruiz, L.A.; Recio, J.A.; Fernández-Sarría, A.; Hermosilla, T. A feature extraction software tool for agricultural object-based image analysis. *Comput. Electron. Agric.* **2011**, *76*, 284–296. [\[CrossRef\]](#)
61. Nilsson, R.; Peña, J.M.; Björkegren, J.; Tegnér, J. Consistent Feature Selection for Pattern Recognition in Polynomial Time. *J. Mach. Learn. Res.* **2007**, *8*, 589–612.
62. Kursa, M.B.; Rudnicki, W.R. R Package 'Boruta'. Available online: <https://cran.r-project.org/web/packages/Boruta/Boruta.pdf> (accessed on 4 August 2018).
63. Deng, H. *Interpreting Tree Ensembles with inTrees*; Springer: Berlin, Germany, 2014.
64. Horrocks, I.; Patel-Schneider, P.F.; Boley, H.; Tabet, S.; Grosz, B.; Dean, M. SWRL: A semantic web rule language combining OWL and RuleML. *W3C Memb. Submiss.* **2004**, *21*, 79.

65. Motik, B.; Patel-Schneider, P.F.; Parsia, B.; Bock, C.; Fokoue, A.; Haase, P.; Hoekstra, R.; Horrocks, I.; Ruttenberg, A.; Sattler, U. OWL 2 web ontology language: Structural specification and functional-style syntax. *W3C Recomm.* **2009**, *27*, 159.
66. Sirin, E.; Parsia, B.; Grau, B.C.; Kalyanpur, A.; Katz, Y. Pellet: A practical OWL-DL reasoner. *Web Semant. Sci. Serv. Agents World Wide Web* **2007**, *5*, 51–53. [[CrossRef](#)]
67. Altmann, A.; Toloşi, L.; Sander, O.; Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **2010**, *26*, 1340–1347. [[CrossRef](#)] [[PubMed](#)]
68. Szymczak, S.; Holzinger, E.; Dasgupta, A.; Malley, J.D.; Molloy, A.M.; Mills, J.L.; Brody, L.C.; Stambolian, D.; Bailey-Wilson, J.E. r2VIM: A new variable selection method for random forests in genome-wide association studies. *BioData Min.* **2016**, *9*, 7. [[CrossRef](#)] [[PubMed](#)]
69. Janitza, S.; Celik, E.; Boulesteix, A.L. A computationally fast variable importance test for random forests for high-dimensional data. *Adv. Data Anal. Classif.* **2016**. doi:10.1007/s11634-016-0270-x. [[CrossRef](#)]
70. Degenhardt, F.; Seifert, S.; Szymczak, S. Evaluation of variable selection methods for random forests and omics data sets. *Brief. Bioinform.* **2017**, doi:10.1093/bib/bbx124. [[CrossRef](#)] [[PubMed](#)]
71. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
72. St-Onge, B.A.; Cavayas, F. Automated forest structure mapping from high resolution imagery based on directional semivariogram estimates. *Remote Sens. Environ.* **1997**, *61*, 82–95. [[CrossRef](#)]
73. Yue, A.; Zhang, C.; Yang, J.; Su, W.; Yun, W.; Zhu, D. Texture extraction for object-oriented classification of high spatial resolution remotely sensed images using a semivariogram. *Int. J. Remote Sens.* **2013**, *34*, 3736–3759. [[CrossRef](#)]
74. Murray, H.; Lucieer, A.; Williams, R. Texture-based classification of sub-Antarctic vegetation communities on Heard Island. *Int. J. Appl. Earth Observ. Geoinf.* **2010**, *12*, 138–149. [[CrossRef](#)]
75. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16.
76. Lang, S. Object-based image analysis for remote sensing applications: Modeling reality—Dealing with complexity. In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*; Blaschke, T., Lang, S., Hay, G.J., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 3–27.
77. Akmal, S.; Shih, L.H.; Batres, R. Ontology-based similarity for product information retrieval. *Comput. Ind.* **2014**, *65*, 91–107. [[CrossRef](#)]
78. Snchez, D.; Moreno, A. Learning non-taxonomic relationships from web documents for domain ontology construction. *Data Knowl. Eng.* **2008**, *64*, 600–623. [[CrossRef](#)]

Sample Availability: The codes developed in R language are available upon request to the corresponding author for testing the replicability of this research.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).